

An aerial night view of a city, likely Shanghai, featuring prominent skyscrapers like the Jin Mao Tower and the Oriental Pearl Tower. The image is overlaid with a network of glowing white lines connecting various points, and several glowing location pin icons are scattered across the scene, suggesting a global or interconnected theme.

LINCOLN

CHALLENGE

Soirée de lancement

Challenge LINCOLN : 6^{ème} édition

*« Je redoute trois jours plus que 100.000 baïonnettes »
Napoléon Bonaparte*

Le challenge Data Science, c'est :

- Participer de manière **collaborative**
- **Monter en compétence** sur des outils, des nouvelles problématiques
- **Partager** et découvrir des nouvelles méthodes analytiques
- **S'amuser** et laisser libre cours à sa **créativité**

...dans une **ambiance conviviale** et **sans pression** !

Challenge News Data Media



Equipes



Thématique

Traitement de journaux et d'émissions Radio



Données

Image de journaux – Textes d'émissions



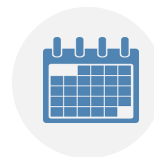
Livrables

Données qualifiées & Présentation



Critères d'évaluation

Performance, Démarche & Restitution



Planning



13 équipes



Tiphaine VAULEON
Claire UTIEL



Moctar EL MAHJOUB
Juanita DAGNON



Louis-Alexandre BAYOL
Oualid AITBOUHOU
Martin FELDMAN



Liang LIANG
Dorian FREMONT



Lydia VIJAYARAJAH
Jeanne CELERIER



Fayçal BENSaid
Abderrahim TALEB



Thomas NICOLAS
Julien-Yacine CHAQRA



Steven MACAIGNE



Antoine HAMON



Axel-Philippe GUEHI



Jack SWINDLE



Antoine VOKLEBER



Théo DAVILLE



Objectif du challenge

- Traiter un sujet lié à l'actualité à partir des sources de données proposées
 - ❖ Images PDF de quotidiens sur le mois de septembre 2019
 - CNEWS
 - 20 minutes
 - ❖ Transcription d'émissions Radio sur le mois de septembre 2019
 - RTL
 - France inter
 - Europe 1





Quelques propositions de sujets



20
minutes

C NEWS



Europe 1

RTL

france
inter

- Segmentation de la page afin d'extraire la structure et l'information du journal
- Détection et description des publicités
- Comparaison du traitement d'un sujet d'actualité entre deux journaux
- Détection et résolution automatique de jeux

- Création de résumé de l'actualité
- Détection des personnes et de leur rôle dans l'émission
- Comparaison du traitement d'une information entre émissions de Radio

Vous êtes libre de créer votre propre sujet !
Seule contrainte : utiliser les données fournies



Extraction de la structure et de l'information

SPORTS

Ce quelque chose de dionysien...

Football Andorre, 136^e nation mondiale, aura droit aux honneurs du Stade de France (20h45), que les Bleus ne quittent décidément plus

Les amicaux pour la province, et les gros matchs qualificatifs pour la capitale. Ce mardi (20h45), les supporters seront 50 000 à remplir les tribunes du Stade de France pour voir l'équipe de France affronter la 136^e nation mondiale dans le cadre des éliminatoires de l'Euro 2020. Un match que le sélectionneur Didier Deschamps, n'a jamais envisagé ailleurs qu'à Saint-Denis. D'autant plus que ce sera le centième joué par les Bleus dans l'enceinte dionysienne.

« Des ondes positives »

Une chose est sûre, l'équipe de France a du mal à quitter son cocon. Un tournant, celui d'une génération peut-être : France-Ukraine, en 2015. « Cette rencontre (3-0 après avoir perdu l'aller 2-0) a marqué les mémoires pour qui était présent ce jour-là, se souvient le portier Hugo Lloris. Le Stade de France

fait partie du patrimoine, on l'associe plutôt à des ondes positives. » Ainsi, les Bleus continuent à organiser les rencontres qualificatives à un tournoi majeur – même celles qui ne font pas rêver – au Stade de France. Et pour cause. Luc Holtz, l'homologue luxembourgeois de Deschamps, avait témoigné de sa colère, en septembre 2017, de devoir affronter les Bleus à Toulouse plutôt qu'à Saint-Denis. Il y trouvait même un écho pour transcrire ses hommes lors de la cause d'avant-match, sur l'air : « On est tellement des ploucs pour eux qu'ils nous jouent dans un stade de kermesse ». Résultats : 0-0 et « DD » qui décide de mettre fin aux tournées dans l'Hexagone. L'intérêt sportif avant tout. « Le Stade de France paraît plus grand que les autres stades, comme s'il y avait plus d'espace, assure Deschamps. Ça joue



Cinquante mille fans sont attendus pour la réception d'Andorre, ce mardi.

Une équipe remaniée

Didier Deschamps a assuré, en conférence de presse mardi, qu'il effectuerait des changements dans son onze de départ par rapport au match face à l'Albanie, samedi 14-11. « Cela me semble logique, a assuré le sélectionneur des Bleus. Ce ne sera pas la même équipe, c'est certain. » Le technicien a assuré que le défenseur Lucas Hernandez, qui avait reçu un coup lors de France-Albanie, était apte pour le rendez-vous face à l'Andorre.

Julien Lalo

10
Mardi 10 septembre 2019

La France perd sur le fil et défilera les Etats-Unis

Basket

Après leur défaite lundi contre l'Australie (98-100), les Français n'échappent pas au choc face aux Etats-Unis en quart de finale du Mondial chinois. « On ne pouvait pas perdre en marquant 98 points avec la défense qu'on avait proposée jusque-là », a analysé après le match le capitaine, Nicolas Batum. Mais les Bleus n'ont pas su cette fois freiner les offensives adverses. Même son de cloche du côté du sélectionneur, qui relève le positif. « C'est ça qui crève le cœur, on a fait un match incroyable, presque au-delà de ce que je pouvais espérer dans le domaine offensif », a confié Vincent Collet.

Et du positif, il en faudra pour les Bleus avant de défier les Etats-Unis, mercredi (13h). La team US a beau avoir emmené une équipe C ou D en Chine, sans aucune de ses grandes stars, elle aligne tout de même 12 joueurs de NBA que tous donnent favoris pour l'or. Mais tout espoir n'est pas éteint. « C'est dur parce qu'on n'aime pas perdre, mais on n'est pas morts », a assuré Nicolas Batum.



L'Espagnol compte 19 victoires en tournoi du Grand Chelem.

Nadal ne devrait pas se faire prier pour éclipser l'étoile Federer

Tennis

20-19 en 2019, 20-20 en 2020 ? Toujours derrière Federer au nombre de tournois du Grand Chelem gagnés, Rafael Nadal a fait un pas en avant en s'adjugeant la nuit dernière, l'US Open au détriment du Russe Daniil Medvedev (7-5, 6-3, 5-7, 4-6, 6-4). Le début de la fin pour « Rodgeur » ?

A Flushing Meadows, les blessures ont rattrapé Federer et Djokovic. « Rafa », lui, semble être revenu à son meilleur niveau, sans jamais perdre son envie de gagner. « La manière dont le match a pris une tournure dramatique à la fin rend ce jour inoubliable pour moi, a expliqué Nadal après son sacre américain. Ce trophée est tellement important pour moi. » Et le Taureau de

« Rafa » pourrait détrôner « Rodgeur » lors du prochain Roland-Garros.

Manacor, âgé de 33 ans, devrait être en mesure de dompter la chaleur et l'humidité de Melbourne pour être à égalité avec le Suisse des Australiens en début de saison. Ensuite, pas la peine d'épiloguer sur le résultat à Roland-Garros, où Nadal gagnera son 21^e tournoi du Grand Chelem. A Federer enfin de remporter une neuvième fois le tournoi de Wimbledon pour redonner du suspense à cette course sans ligne d'arrivée. A.H.

20 secondes

Sacré couac pour les Croates.

Lundi, la Croatie, finaliste du Mondial en 2018, a été accrochée à la surprise générale par l'Azerbaïdjan (1-1), qui marque son premier point dans la course à la qualification pour l'Euro de football 2020. Dans son groupe, l'équipe de Luka Modric va devoir batailler notamment avec la Hongrie et la Slovaquie.

Napeth ménagé au début de l'Eurovolley.

Lundi, Earvin Ngapeth a passé une IRM révélant une lésion au niveau des côtes, rapporte l'Equipe. La star de l'équipe de France de volley manquera les trois premiers matchs du championnat d'Europe contre la Roumanie, jeudi, la Grèce, samedi, et le Portugal, dimanche.

Fuglsang vainqueur, Roglic solide leader de la Vuelta.

Le Danois Jakob Fuglsang (Astana) a remporté, lundi, la 16^e étape du Tour d'Espagne au sommet de La Cubilla. Un temps mis en difficulté, le maillot rouge Primoz Roglic (Jumbo-Visma) a réussi à accentuer son avance au général.

Ce quelque chose de dionysien...

Football Andorre, 136^e nation mondiale, aura droit aux honneurs du Stade de France (20h45), que les Bleus ne quittent décidément plus

les amicaux pour la province, et les gros matchs qualificatifs pour la capitale. Ce mardi (20h45), les supporters seront 50 000 à remplir les tribunes du Stade de France pour voir l'équipe de France affronter la 136^e nation mondiale dans le cadre des éliminatoires de l'Euro 2020. Un match que le sélectionneur Didier Deschamps, n'a jamais envisagé ailleurs qu'à Saint-Denis. D'autant plus que ce sera le centième joué par les Bleus dans l'enceinte dionysienne.

« Des ondes positives »

Une chose est sûre, l'équipe de France a du mal à quitter son cocon. Un tournant, celui d'une génération peut-être : France-Ukraine, en 2015. « Cette rencontre (3-0 après avoir perdu l'aller 2-0) a marqué les mémoires pour qui était présent ce jour-là, se souvient le portier Hugo Lloris. Le Stade de France

fait partie du patrimoine, on l'associe plutôt à des ondes positives. » Ainsi, les Bleus continuent à organiser les rencontres qualificatives à un tournoi majeur – même celles qui ne font pas rêver – au Stade de France. Et pour cause. Luc Holtz, l'homologue luxembourgeois de Deschamps, avait témoigné de sa colère, en septembre 2017, de devoir affronter les Bleus à Toulouse plutôt qu'à Saint-Denis. Il y trouvait même un écho pour transcrire ses hommes lors de la cause d'avant-match, sur l'air : « On est tellement des ploucs pour eux qu'ils nous jouent dans un stade de kermesse ». Résultats : 0-0 et « DD » qui décide de mettre fin aux tournées dans l'Hexagone. L'intérêt sportif avant tout. « Le Stade de France paraît plus grand que les autres stades, comme s'il y avait plus d'espace, assure Deschamps. Ça joue



Cinquante mille fans sont attendus pour la réception d'Andorre, ce mardi.

Une équipe remaniée

Didier Deschamps a assuré, en conférence de presse mardi, qu'il effectuerait des changements dans son onze de départ par rapport au match face à l'Albanie, samedi 14-11. « Cela me semble logique, a assuré le sélectionneur des Bleus. Ce ne sera pas la même équipe, c'est certain. » Le technicien a assuré que le défenseur Lucas Hernandez, qui avait reçu un coup lors de France-Albanie, était apte pour le rendez-vous face à l'Andorre.

Julien Lalo

Article



Cinquante mille fans sont attendus pour la réception d'Andorre, ce mardi.

Image

20 secondes

Sacré couac pour les Croates.

Lundi, la Croatie, finaliste du Mondial en 2018, a été accrochée à la surprise générale par l'Azerbaïdjan (1-1), qui marque son premier point dans la course à la qualification pour l'Euro de football 2020. Dans son groupe, l'équipe de Luka Modric va devoir batailler notamment avec la Hongrie et la Slovaquie.

Napeth ménagé au début de l'Eurovolley.

Lundi, Earvin Ngapeth a passé une IRM révélant une lésion au niveau des côtes, rapporte l'Equipe. La star de l'équipe de France de volley manquera les trois premiers matchs du championnat d'Europe contre la Roumanie, jeudi, la Grèce, samedi, et le Portugal, dimanche.

Fuglsang vainqueur, Roglic solide leader de la Vuelta.

Le Danois Jakob Fuglsang (Astana) a remporté, lundi, la 16^e étape du Tour d'Espagne au sommet de La Cubilla. Un temps mis en difficulté, le maillot rouge Primoz Roglic (Jumbo-Visma) a réussi à accentuer son avance au général.

Sacré couac pour les Croates.

Lundi, la Croatie, finaliste du Mondial en 2018, a été accrochée à la surprise générale par l'Azerbaïdjan (1-1), qui marque son premier point dans la course à la qualification pour l'Euro de football 2020. Dans son groupe, l'équipe de Luka Modric va devoir batailler notamment avec la Hongrie et la Slovaquie.

Article

Page

Rubrique

LINCOLN



Extraction de la structure et de l'information

Thématiques

- Document Layout Analysis
- Zone segmentation
- Texture segmentation
- Optical Character Recognition (OCR)

Champs d'application

- Extraction de données
- Création de modèles de données

Ressources

- Multitask handwritten document layout analysis :
 - <https://arxiv.org/pdf/1806.08852v3.pdf>
 - <https://github.com/lquiroso/P2PaLA>
- OCR :
 - Tesseract, google vision ...
- Segmentation de page par recouvrement des espaces blancs :
 - <https://hal.inria.fr/hal-00779235/document>
- Texture segmentation :
 - <http://arima.inria.fr/006/pdf/arima00604.pdf>



Détection et description des publicités

ACTUALITÉ

Un droit fécond en interrogations

Bioéthique Manque de donneurs, de moyens... La procréation médicalement assistée (PMA) pour toutes pose certains défis pratiques

On parle beaucoup FM, don de sperme et égalité des droits en cette rentrée. En effet, les auditions à la commission spéciale chargée d'examiner le projet de loi sur la bioéthique, qui prévoit notamment de donner la possibilité aux couples de lesbiennes et aux femmes seules de recourir à une procréation médicalement assistée (PMA) pour avoir un enfant, ont commencé mardi. Une ci-dessous. Mais, déjà, certaines associations et certains médecins veulent faire pression pour que le texte ne passe pas à côté d'un point essentiel : les conditions matérielles pour garantir que cette PMA pour toutes soit réellement assurée. Sans surprise, l'ouverture de la PMA à toutes les femmes augmenterait la demande de sperme. Or, aujourd'hui, les couples hétérosexuels qui ont recours à un don de sperme - 2200 en 2016 - attendent généralement entre six mois et

deux ans. Ce délai risque donc de s'allonger. D'autant que le projet de loi de bioéthique introduit plusieurs changements, au-delà de la PMA pour toutes. Notamment la possibilité pour les enfants nés d'une PMA avec don d'avoir accès, s'ils le souhaitent et à leur majorité, à l'identité du donneur. Une disposition qui pourrait en refroidir certains.

Stocks de sperme détruits ? Deuxième problème, le texte prévoit la destruction du stock de gamètes actuel un an après la promulgation de la loi pour éviter la coexistence des dons identifiants et de ceux qui ne le sont pas. « Être obligé de détruire [les stocks], c'est de la folie furieuse », s'élève Catherine Guillemin, vice-présidente de la Fédération des Cecos, les centres qui gèrent les dons de gamètes. La seule possibilité, pour sauver ces

paillottes, serait de recontacter tous les donneurs. Mais au cabinet de la ministre de la Santé, les choses sont claires : « Il est hors de question de recontacter tous les donneurs. C'est matériellement impossible, certains ont pu déménager, par exemple. On peut en revanche envisager d'appeler publiquement les anciens donneurs à se manifester. » Du côté des députés, les auditions devraient permettre d'affiner les choses. « Nous souhaitons entendre la reconstitution de stocks équilibrés de gamètes avant de détruire les anciennes réserves », assure Aurélie Bergé, députée LREM et membre de la



Les délais pour obtenir un don de sperme vont de six mois à deux ans.

commission spéciale bioéthique. Mais quand bien même. « Sans changement et moyens supplémentaires, on aura du mal à répondre à la demande actuelle et à celle qui se profile », prévient Catherine Guillemin. Face à ces défis, 350 professionnels de la PMA demandent via une pétition que la loi permette à d'autres centres privés, de gérer les dons de gamètes et l'autocconservation - ces derniers le sont actuellement par les seuls Cecos, au nombre de 29. Cela réglerait-il pour autant la question du nombre de dons ?

Oliana Gabriel

6
Vendredi 30 août 2019

1. Extraction de la publicité



2. Description de la publicité

- Publicité
- Masha et Michka
- Ours et enfant qui tient une loupe
- 10 dernières représentations
- ...



« Le niveau de connaissances des gens sur ces questions s'est accru »

Mardi, les collectifs pro et anti-PMA ont ouvert le bal des auditions à l'Assemblée. Jean-Louis Touraine, médecin, député LREM et rapporteur du texte, apporte son éclairage sur le débat.

Que ressort-il des premières auditions ? Je vois que le niveau de connaissances des gens sur les questions de bioéthique s'est accru. Et je remarque une adhésion large au principe de l'extension de la PMA, non seulement chez LREM, à gauche, mais même à droite.

Dans Donner la vie, choisir sa mort. Pour une bioéthique de liberté (P 25), vous évoquez la question du bien-être de ces enfants élevés sans père. Que disent les études ? Certains reprochent au projet de loi de faire des familles sans père. C'est vrai. Est-ce un handicap ? Toutes les études américaines montrent que le développement des enfants nés de PMA dans une famille hétérosexuelle, homosexuelle, de femmes seules, n'offre aucune différence avec celui des enfants nés de façon « naturelle ». De plus, quand on interroge ces enfants nés dans des couples de femmes, ils disent : « Un père, pour



Le rapporteur Jean-Louis Touraine.



Détection et description des publicités

Thématiques

- Document Layout Analysis
- Optical Character Recognition (OCR)
- Clustering
- Scene Parsing / Image captioning

Champs d'application

- Extraction de données
- Analyse textuelle
- Génération automatique de données

Ressources

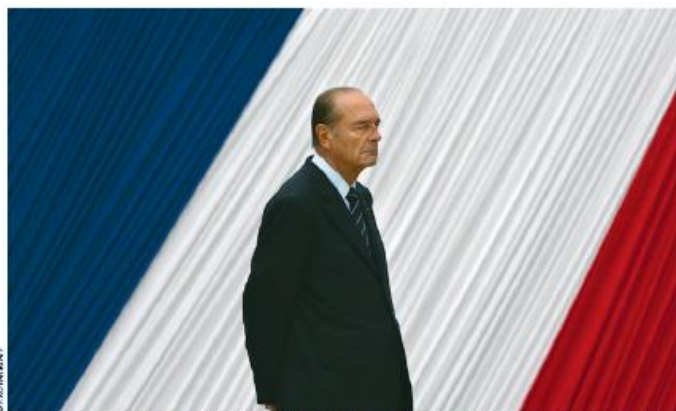
- Scene parsing / Semantic segmentation
 - <https://github.com/CSAILVision/semantic-segmentation-pytorch>
 - https://www.tensorflow.org/tutorials/text/image_captioning
- OCR : Tesseract, google vision ...
- Text clustering : <http://charuaggarwal.net/text-cluster.pdf>
- Embeddings : https://fr.wikipedia.org/wiki/Word_embedding



Comparaison du traitement d'un sujet d'actualité entre 2 journaux

L'ANCIEN PRÉSIDENT JACQUES CHIRAC EST DÉCÉDÉ À L'ÂGE DE 86 ANS

LA FRANCE DANS LE CŒUR



Lors de son passage à l'Élysée, le fondateur du RPR a notamment refusé de s'impliquer dans la guerre en Irak.

L'homme d'Etat laissera une empreinte indélébile sur la politique de son pays, tout en restant extrêmement populaire.

Une figure de la V^e République.

Jacques Chirac, président de 1995 à 2007, est décédé à l'âge de 86 ans, a annoncé hier sa famille. Pluie d'hommages venant de personnalités de tous bords, minute de silence à l'Assemblée et au Sénat, drapeaux en berne à Paris, deuil national lundi. Hier soir, Emmanuel Macron a tenu à rappeler « l'amour de la France et des français » que portait celui qui « occupa les plus hautes fonctions sans jamais oublier ses racines ». Cette nouvelle a provoqué d'autres réactions émus, illustrant l'impact qu'aura eu « Chirac » sur tous, durant ses douze années à la tête du pays, et plus de quarante à différents postes de pouvoir.

De la Corrèze à l'Élysée

De ses premiers pas en politique en tant que conseiller municipal et député de Corrèze, en 1967, jusqu'à son accession à l'Élysée, en 1995, Jacques Chirac a tout connu. Secrétaire d'Etat et ministre à plusieurs reprises sous Charles de Gaulle et Georges Pompidou, cet ancien de l'ENA a été nommé Premier ministre par Valéry Giscard d'Estaing en 1974. Preuve de son tempérament, il n'a pas hésité à démissionner en 1976, ne s'entendant pas avec VGE. C'est alors qu'il a fondé le RPR (devenu UMP, puis Les Républicains), et s'est lancé à la conquête de la mairie de Paris, qu'il a dirigée de 1977 à 1995. Durant cette

période, il est devenu, de 1986 à 1988, le premier chef d'un gouvernement de cohabitation, sous François Mitterrand. S'il a échoué à s'imposer face au chef d'Etat socialiste en 1988, ce revers n'a fait que renforcer sa détermination. En 1995, il s'est ainsi lancé corps et âme dans la course à la présidentielle malgré la candidature de son « ami » du RPR, Edouard Balladur, et a accédé enfin à la fonction suprême en battant le socialiste Lionel Jospin. Dès son entrée à l'Élysée, Jacques Chirac a été au cœur d'une polémique, en annonçant la relance des essais nucléaires français. Mais il a aussi été l'homme qui a mis fin au septennat présidentiel, en 2000, et qui a incarné le rempart contre le front national de Jean-Marie Le Pen durant la présidentielle 2002. Quant à son second mandat, il a été marqué par son refus de s'impliquer dans la guerre en Irak lancée par le président américain George W. Bush, en 2003.

BIO EXPRESS

29 NOVEMBRE 1932. Jacques Chirac naît à Paris. Enfant unique, il intègre Sciences-Po, avant d'épouser Bernadette, dont il aura deux filles.

1967. Après un premier mandat de conseiller l'énergie devient député de Corrèze. Il rejoint dans la foulée le gouvernement Pompidou.

1974. Il démissionne de Matignon, fonde le RPR et, trois ans plus tard, devient maire de Paris. Il redeviendra Premier ministre en 1986.

1995. Il est élu président de la République, avec 52 % des voix, contre le candidat socialiste Lionel Jospin.

2002. Il est réélu à 82 % des voix contre Jean-Marie Le Pen, après l'élimination surprise du même Lionel Jospin au premier tour.

Nostalgie et sympathie

Mais il n'aura jamais été aussi apprécié que depuis son départ de l'Élysée, en 2007. Ses multiples apparitions au Salon de l'Agriculture, son style, ses phrases cutes, ont en effet marqué plusieurs générations. A tel point que chacune de ses sorties publiques, en Corrèze ou à Saint-Tropez, inspirent nostalgie et sympathie. Victime d'un AVC en 2005, ses soucis de santé à partir des années 2010 avaient soulevé des inquiétudes. Mais malgré son absence publique, la France ne l'avait jamais oublié.

L'INFO EN PLUS

Lorsqu'il a démissionné de son poste de Premier ministre en 1976, Jacques Chirac a assuré qu'il « était la politique pour ouvrir une galerie d'art ».

VS

GRAND PARIS

La capitale « en deuil »

Hommage

Dans l'ancien QG de Jacques Chirac, on se souvient de l'homme, du maire, des affaires

La nouvelle a cueilli les larmes, jeudi matin, en pleine conférence de presse en amont du Conseil de Paris. Jacques Chirac, acteur omniprésent de quatre décennies de la vie politique française et internationale, est mort à son domicile parisien à l'âge de 86 ans. L'Hôtel de Ville, où le premier maire de la capitale a pris ses quartiers en 1977, se met soudain en branle. Entre larmes, souvenirs et moments de baillonnement. Les élus de la droite parisienne se succèdent au micro : en tête, Marie-Claire Camille-Gèle, fraîchement élue présidente du groupe LR (Les Républicains et Indépendants). Cette ancienne secrétaire générale adjointe de l'Élysée sous Chirac et conseillère quand il était maire rend spontanément hommage à cette « figure paternelle [...] c'est à Paris que tout a commencé ». « Paris pleure son maire », annonce le groupe



En 1977, au soir de son élection.

dans la foule. Des gens courent dans les couloirs, les hommages pleuvent : Bertrand Delanoë, Jean Tiberi et bien sûr, à six mois des élections municipales, ceux de tous les prétendants au poste : Benjamin Griveaux, Clément Villard, Gaëtan Gantier.

Dans les rangs de la gauche parisienne, on est en revanche moins ému. « Les amendes valent toujours en excédent », avait déclaré un jour Jacques Chirac. Car Chirac, maire de Paris, ce n'est pas seulement

l'apogée de la célèbre métacritique et la propriété de la ville en thème de campagne. C'est aussi de nombreux scandales politico-financiers, emplois fictifs, faux chèques, affaires des HLM. Pour Lika Filocha, coprésidente du groupe Génération.s, Chirac, le chef d'Etat, est même « à l'origine de son engagement militant et politique » avec l'expulsion des sans-papiers de l'église Saint-Ambroise, en 1996.

La tour Eiffel étoilée

Dans le bureau historique des maires, un portrait de Jacques Chirac, en noir et blanc. Les drapeaux sont en berne. « Paris est en deuil », déclare Anne Hidalgo. « Pour nous les Parisiennes et les Parisiens, il sera à jamais notre maire, aimant passionnément sa ville et ses habitants », a-t-elle ajouté sur Twitter. En hommage, les drapeaux des équipements municipaux ont été mis en berne et la tour Eiffel, étoilée jeudi soir. Des registres seront ouverts à l'Hôtel de Ville et dans toutes les mairies d'arrondissement « pour permettre aux Parisiennes et aux Parisiens de témoigner leur affection », a annoncé Anne Hidalgo.

Romain Lecomte



Comparaison du traitement d'un sujet d'actualité entre 2 journaux

Thématiques

- Sentiment analysis
- Text classification
- Morphological analysis
- Topic modeling

Champs d'application

- Traitement de bases documentaires
- Etudes de verbatims
- Détection de thématiques abordées

Ressources

- Comparative analysis of newspapers : <https://www.aclweb.org/anthology/Y13-1047.pdf>
- OCR : Tesseract, google vision ...
- Latent Dirichlet allocation : https://rstudio-pubs-static.s3.amazonaws.com/79360_850b2a69980c4488b1db95987a24867a.html
- Embeddings : https://fr.wikipedia.org/wiki/Word_embedding



Résolution automatique de jeux

PAUSE

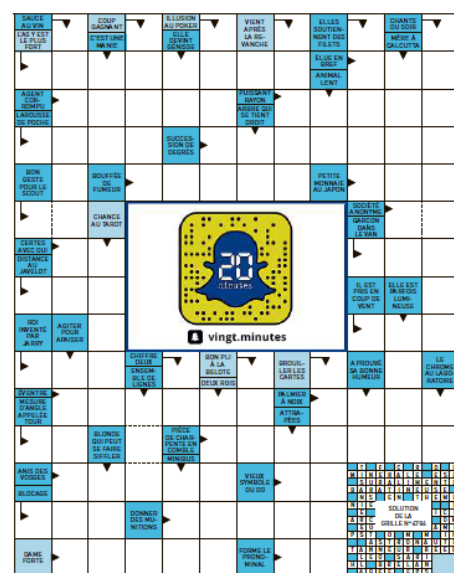
Horoscope

- Bélier du 21 mars au 20 avril**
Si vos idées surprennent vos collègues, exposez vos projets à vos supérieurs.
- Taureau du 21 avril au 21 mai**
Des nuages obscurcissent votre ciel, mais ils se dissipent assez vite. Donnez-vous un objectif avant tout.
- Gémeaux du 22 mai au 21 juin**
Ça coince un peu avec l'âme sœur. Ne boudez pas, renouez le dialogue. Des dépenses inattendues sont à prévoir.
- Cancer du 22 juin au 22 juillet**
Un défi vous attend, soyez à la hauteur. Vous mettez en avant votre créativité tout en allant à l'essentiel.
- Lion du 23 juillet au 23 août**
N'hésitez pas à déclarer votre flamme à la personne dont vous êtes épris. Retrouvez vos manches et au travail !
- Vierge du 24 août au 23 septembre**
Vous aimez faire plaisir à vos proches, mais ne promettez rien d'impossible. Vos projets vous prennent du temps.
- Balance du 24 sept. au 23 octobre**
L'ambiance familiale est tendue. Il faut dire que vous êtes irritable. La spéculation financière est payante.
- Scorpion du 24 oct. au 22 nov.**
Vous êtes sûr de vous, mais réfléchissez tout de même avant de rater une belle opportunité.
- Sagittaire du 23 nov. au 21 déc.**
Souplesse et adaptabilité, deux qualités dont vous allez devoir faire preuve aujourd'hui.
- Capricorne du 22 déc. au 20 janv.**
Aucun obstacle ne s'oppose à ce que vous réalisiez une affaire. Mais faites preuve de plus de conviction.
- Verseau du 21 janv. au 18 février**
Vous seul savez écarier les embûches qui vous éloignent de l'être aimé. Vos tâches du jour devraient vous ravir.
- Poissons du 19 février au 20 mars**
En couple, une vendue complotée régnait. Il vous faut expliquer plus clairement votre point de vue à vos associés.
- Cerf-panthère pas avant midi**
C'est vachement long la rentrée cette année, non ?

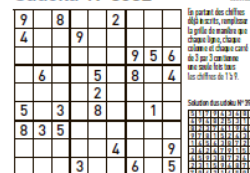
Retrouvez l'horoscope complet sur le site de nos apps.

20 minutes
3^e marque de presse française
22,1 millions d'abonnés par mois
1^{er} quotidien avec 3000000 d'abonnés
Actionnaires : Sofina, Bouygues
NORMONE (Gazet 2015) (LNE 2017)
24-25 rue du Centre, 92210 St-Denis
01 53 26 65 65 - Fax 01 53 26 65 10

Mots fléchés N° 4783



Sudoku N° 3952



17
Vendredi 30 août 2019

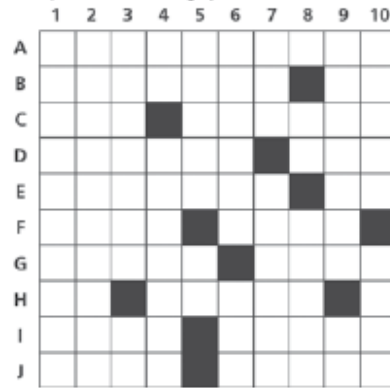
Jeux de cartes

MOTS CROISES

HORIZONTALEMENT A- Organisation et gestion d'une entreprise. B- Match nul. Fin de forme verbale. C- Démonstratif. Celui du loto est diffusé sur le petit écran. D- Jouer le boute-en-train. Service de réparations. E- Louer un navire marchand. Le pilote et l'architecte en ont l'utilité. F- S'il n'est pas chaton, il s'enroule facilement. Personnage de crèche provençale. G- Baldaquins. Il suit le dessert. H- Sujet pour homme. Son coup fait mal au cou. I- Choisie. Elle débute par une fête. J- Neige persistante. Effectifs.

VERTICALEMENT

- 1- Professionnel du moulin, mais pas meunier. 2- En position pour prier. 3- Parle du nez. Rayons dont il ne faut pas abuser. 4- L'aluminium du chimiste. De couleur ambrée et de consistance sirupeuse et sucrée. 5- Abris pour hases et levrauts. Possessif. 6- Augmenter les dimensions. Score idéal pour un golfeur. 7- Réserve naturelle de sel. Elle se développe sous terre. 8- Carte maîtresse. Son ouverture et sa fermeture règle le débit de l'eau. 9- Epreuve photographique. Article étranger. 10- Cessez-le-feu. Façon d'être.





Résolution automatique de jeux

Thématiques

- OCR
- Grid extraction
- Solveur de jeux

Champs d'application

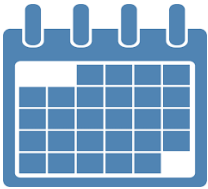
- Extraction d'information
- Algorithmique

Ressources

- Extraction de grille
 - <http://honors.cs.umd.edu/uploads/thesis/file/178/Thesis.pdf>
 - <https://github.com/dgulotta/cross2sheet>
 - <https://github.com/HartasCuerdas/xwbinarify>
- Solveur (exemple pour le sudoku)
 - <http://igm.univ-mlv.fr/~dr/XPOSE2013/sudoku/>



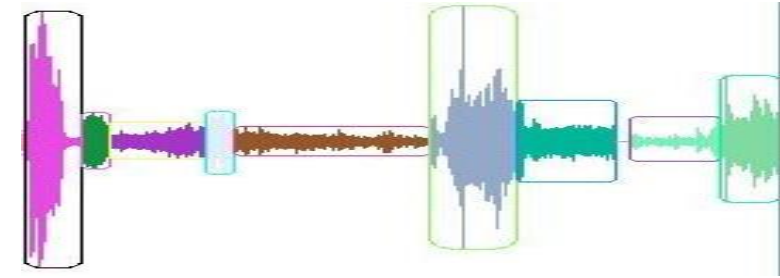
Emissions Radio : les données



Du 2 Septembre 2019
au 26 septembre 2019



L'actu sur la plage horaire 7h-9h



Fichier ***donnees_radio.csv*** contenant les émissions découpées en segments avec les colonnes suivantes :

- **debut_seg** : temps de début du segment (cs)
- **duree_seg** : durée du segment (cs)
- **genre** : genre du locuteur (M ou F)
- **locuteur** : identifiant du locuteur (par émission/jour)
- **texte** : texte transcrit (sans ponctuation)
- **date** : date de l'émission (AAAA-MM-JJ)
- **emission** : émission



Emissions Radio : les données

Un fichier ***donnees_radio_thematisees.csv*** contenant les émissions découpées en segments / groupes de souffle / mots lemmatisés. Ce fichier contient les colonnes suivantes :

- **id_emission** : concaténation de la chaîne Radio et de la date
- **id_segment** : Identifiant du segment
- **mot** : mot
- **id_GS**: identifiant du groupe de souffle
- **cluster** : cluster thématique auquel le mot appartient (tous les mots d'un groupe de souffle appartiennent au même cluster)

Note : un groupe de souffle est un ensemble de mots prononcés sans prendre de pause de respiration.



Résumé de l'actualité

Thématiques

- Text summarization
- Topic modeling

Champs d'application

- Exploration de documents
- Traitement de données textuelles
- Visualisation de texte

Ressources

- Text summarization
 - <https://arxiv.org/pdf/1801.10198v1.pdf>
 - <https://www.aclweb.org/anthology/W17-1003.pdf>
- Evaluation : <https://rxnlp.com/how-rouge-works-for-evaluation-of-summarization-tasks/>
- <https://github.com/miso-belica/sumy>
- <https://github.com/mathsyouth/awesome-text-summarization>
- <https://arxiv.org/ftp/arxiv/papers/1304/1304.2476.pdf>



Identification des locuteurs

Thématiques

- Détection des entités nommées
- Identification des rôles/rerelations

Champs d'application

- Extraction d'informations
- Structuration de l'information
- Visualisation de relations

Ressources

- https://fr.wikipedia.org/wiki/Reconnaissance_d%27entit%C3%A9s_nomm%C3%A9es
- <https://github.com/JuneFeng/RelationClassification-RL>
- <https://arxiv.org/pdf/1904.10887.pdf>



Comparaison du traitement d'un sujet d'actualité

Thématiques

- Détection de thématiques / sujets traités
- Analyse de sentiments
- Word embedding

Champs d'application

- Exploration de base documentaire
- Etude de verbatims
- Détection de thématiques abordées

Ressources

- Détection de thématiques : <https://openclassrooms.com/fr/courses/4470541-analysez-vos-donnees-textuelles/4855011-modelisez-des-sujets-avec-des-methodes-non-supervisees>
- Word embedding :
 - <https://openclassrooms.com/fr/courses/4470541-analysez-vos-donnees-textuelles/4855006-effectuez-des-plongements-de-mots-word-embeddings>
 - https://fr.wikipedia.org/wiki/Word_embedding
- Analyse de sentiment : <https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17>



Livrables

- Un **document PowerPoint de 15 slides maximum** présentant l'objectif fixé, la démarche méthodologique, les traitements réalisés, les résultats obtenus et commentés et **des scénarios d'application**.
- **Le code**, qui pourra être partagé sur Gitea

Optionnel :

- **Une vidéo** ou tout support animé de trois minutes maximum



Evaluation

- Pour évaluer les restitutions, le jury se basera sur les **critères suivants** :
 - ❖ Originalité et pertinence de l'approche
Ce critère juge de l'originalité de la démarche retenue et de la pertinence des traitements effectués
 - ❖ Expertise technique & méthodologique
Ce critère vise à mesurer les techniques utilisées pour traiter la problématique
 - ❖ Qualité des livrables
Ce critère évalue la qualité du code et de la présentation PPT proposée (forme, pédagogie, illustration...)



Composition du jury

- Un jury constitué de **profils complémentaires**
 - ❖ Sydney BECU, Data Scientist
 - ❖ François VIEILLE, Data Scientist
 - ❖ Dorothee DELAUNAY, Responsable de l'Innovation
 - ❖ Maxime BOURSIN, Responsable des Offres
 - ❖ Edouard STEVENINI, Business Manager



Planning

Les grandes dates

LUN	MAR	MER	JEU	VEN	SAM	DIM
Octobre 2019						
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

LUN	MAR	MER	JEU	VEN	SAM	DIM
Novembre 2019						
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

LUN	MAR	MER	JEU	VEN	SAM	DIM
Décembre 2019						
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					



29/10 Soirée de lancement



30/10 **Envoi des données** et des consignes



04/11 Remontée du sujet choisi par les équipes



19/11 Atelier d'avancement (*optionnel*)



02/12 **Fin du challenge** : remise des livrables



06/12 Annonce des finalistes



10/12 Soirée de clôture : présentation & remise des prix



Soirée de clôture

- **06 décembre** : le jury annoncera la **liste des équipes** qui seront amenées à **présenter leurs travaux lors de la soirée de clôture**

- **Programme du 10 décembre**
 - Bilan par l'équipe organisatrice
 - Tour de table des équipes : 2 min par équipe pour présenter sa démarche
 - Présentation par les équipes finalistes : 10 minutes par équipe
 - Podium et remise des prix
 - Cocktail dînatoire

A l'issue du challenge, **toutes les restitutions seront partagées à l'ensemble des participants**

LINCOLN

CHALLENGE

A hand is visible on the left side of the image, pointing towards the first wireframe lightbulb. The background is a dark blue gradient with a row of five wireframe lightbulbs hanging from thin white lines. The lightbulbs are made of a white geometric mesh. The first lightbulb is slightly larger and more prominent than the others.

Merci à tous

A vous de jouer !